

ANALISIS KLASIFIKASI SUARA BERDASARKAN GENDER DENGAN FORMAT WAV MENGUNAKAN ALGORITMA K-MEANS

Arif Setiawan¹

email : arif_plstuff@yahoo.com

arif2301@gmail.com

ABSTRAK

Klasifikasi suara adalah salah satu bidang dari Artificial Intelligent yaitu pada bidang ilmu Patern Recognition, analisis klasifikasi suara berdasarkan gender dengan format wav menggunakan algoritma K-Means adalah sistem cerdas yang dibuat agar komputer dapat membedakan suara pria dan wanita melalui sistem.

Metode yang digunakan dalam penelitian ini adalah menggunakan ekstraksi ciri yang terdiri dari : Sort Time Energy, Zero Crossing Rate, Spectral Centroid, Spectral Flux

Dari penelitian ini, hasil yang didapat adalah pengklasifikasian suara berdasarkan jenis kelamin atau gender dapat dilakukan dengan suatu metode ekstraksi ciri sinyal suara berbasis domain waktu dan domain frekuensi. Nilai rata-rata dari masing-masing ciri dihitung dengan standart deviasi, untuk memperoleh rata-rata nya, lalu diolah untuk pengklasifikasian.

Kata kunci : *Klasifikasi Suara, Wav, Sort Time Energy, Zero Crossing Rate, Spectral Centroid, Spectral Flux*

ABSTRACT

Voice classification is one of the areas of Artificial Intelligent in the field of pattern recognition, classification analysis of gender-based sound with wav format using K-Means algorithm is a clever system that created the computer can distinguish male and female voice through the system.

The method used in this research is to use a feature ekstraksi consisting of: Sort Times Energy, Zero Crossing Rate, Spectral Centroid, Spectral Flux.

From this research, the results obtained are sound classification based on sex or gender can be done with a feature extraction method based voice signal time domain and frequency domain. The average value of each feature is calculated by standard deviation, to obtain his average, and then processed for classification.

Key words : *Voice classification, Wav, Sort Time Energy, Zero Crossing Rate, Spectral Centroid, Spectral Flux*

¹ Dosen Fakultas Teknik Universitas Muria Kudus

1. Pendahuluan

Secara awam, aspek yang cukup penting yang mendasari berbagai teori dalam AI adalah sistem pengenalan pola (*Pattern Recognizing*). Sistem pengenalan pola merupakan komponen penting dalam proses peniruan kemampuan inderawi manusia terutama penglihatan dan pendengaran. Sebagai contoh, untuk meniru indera pendengaran manusia, komputer harus mempunyai suatu mekanisme standar dan logis dalam mengenali pola yang ada pada suara yang sedang diproses. Dari sinilah diperoleh motivasi untuk mencoba suatu konsep sederhana untuk mengenali pola dari suara sehingga dapat diidentifikasi dengan baik oleh komputer.

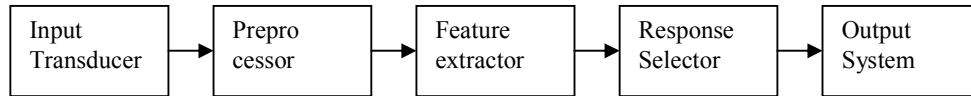
Permasalahan utama yang terjadi apabila hendak mengenali suatu pola tertentu adalah bagaimana proses akuisisi data dilakukan sehingga menghasilkan sejumlah data numerik yang representatif dan konsisten terhadap sampel yang diberikan. Dalam penulisan ini, kita mencoba mengaplikasikan suatu metoda sederhana untuk mengenali suara dan mengklasifikasikannya berdasarkan gender sehingga dapat diidentifikasi dengan baik oleh komputer dengan memanfaatkan berbagai teori ekstraksi ciri untuk data audio

Tujuan utama dari penelitian ini adalah untuk menganalisis dan membuktikan bahwa metoda sederhana untuk mengekstraksi data suara dari sampel yang berupa input suara dari 10 sampel suara, yang terdiri dari 8 pria dan 2 wanita dapat diimplementasikan dengan baik sesuai dengan tujuannya sehingga komputer dapat mengidentifikasi suara secara baik dan konsisten.

2. Pengenalan Pola

Pola adalah entitas yang terdefinisi dan dapat diidentifikasi melalui ciri-cirinya (features). Ciri-ciri tersebut digunakan untuk membedakan suatu pola dengan pola lainnya. Ciri yang bagus adalah ciri yang memiliki daya pembeda yang tinggi, sehingga pengelompokan pola berdasarkan ciri yang dimiliki dapat dilakukan dengan keakuratan yang tinggi.

Pengenalan pola bertujuan menentukan kelompok atau kategori pola berdasarkan ciri-ciri yang dimiliki oleh pola tersebut. Dengan kata lain terdapat dua pendekatan yang dilakukan dalam pengenalan pola, pendekatan secara statistik dan pendekatan secara sintatik atau struktural. Struktur pengenalan pola dapat dilihat dari gambar berikut (gambar 1) :



Gambar 1. Struktur pengenalan pola

a. *Input Transducer*

Menganalisa sinyal elektronik yang diinput. Alat yang digunakan : scanner

b. *Preprocessor*

Melakukan penambahan kondisi sinyal, termasuk fungsi penguatan sinyal, analisa spektrum dan konversi analog ke digital.

c. *Feature Extractor*

Sebagai distriminator, dengan membandingkan kecocokan template.

d. *Response selector*

Algoritma untuk memilih input pattern dengan data pattern yang sudah ada di software dengan menggunakan teknik pencarian, sortir, least-squares analysis.

e. *Output System*

Output system mungkin berupa suara, gambar atau video yang sesuai dengan hasil proses aplikasi.

3. Ekstraksi Ciri

Untuk mendapatkan data yang akurat dan konsisten dari setiap sampel, digunakan suatu metode ekstraksi ciri sinyal suara dengan domain waktu dan domain frekuensi. Pada tugas ini penulis hanya menggunakan masing- masing dua metode untuk tiap domain. Adapun metode yang digunakan adalah sebagai berikut :

1. Time Domain

Time domain adalah pemaparan teknik sinyal audio dasar, dimana sinyal digambarkan sebagai amplitudo dengan satuan waktu, sinyal dapat bernilai positif atau negatif tergantung pada tekanan suara.

a. Sort Time Energy

Menandakan kekerasan suara pada waktu yang pendek

$$STE = \frac{1}{N} \sum_{n=1}^N x(n)^2$$

STE = Sort time energy

N = Jumlah Sampel

X(n) = Nilai sinyal dari sampel

b. Zero Crossing Rate

Sampel berurutan pada sebuah sinyal digital memiliki perbedaan tanda, ukuran dari noise sebuah sinyal pada fitur domain

$$ZC = \frac{\sum_{n=1}^N |\operatorname{sgn} x(n) - \operatorname{sgn} x(n-1)|}{2N}$$

ZC = Zero crossing rate

$\operatorname{sgn} x(n)$ = nilai dari $x(n)$, bernilai 1 jika $x(n)$ positif, -1 jika $x(n)$ negatif

N = Jumlah Sampel

2. Frekuensi Domain

a. Spectral Centroid

Menyeimbangkan titik ukuran spectrum dari bentuk asosiasi spectral dengan spectral brightness. Nilai centroid yang tinggi menunjukkan frekuensi yang tinggi

$$C = \frac{\sum_{n=1}^N Mt(n) \cdot n}{\sum_{n=1}^N Mt(n)}$$

C = Spectral Centroid

Mt = Nilai Centroid

b. Spectral Flux

Ukuran perubahan puncak spectral pada sinyal. Ini didapatkan dari perhitungan perubahan spektrum yang besar dari layer ke layer

$$F = \sum_{n=1}^N (Nt(n) - Nt_{-1}(n))^2$$

F = Spectral Flux

Nt = Nilai perubahan spectrum

Setiap metode ciri diambil rata-rata nya, menggunakan standart deviasi, berikut adalah tabel penggunaan rata-rata untuk tiap-tiap metode ciri.

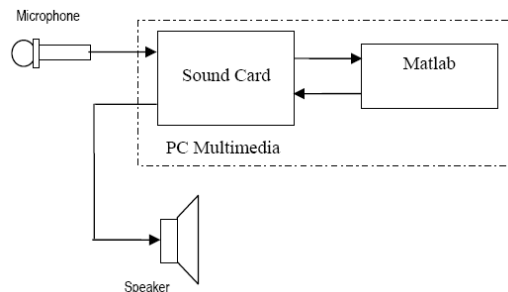
Tabel 1. penggunaan rata-rata untuk tiap-tiap metode ciri.

Ciri	Statistik
Sort Time Energy	Standart Deviasi (std)
Zero Crossing Rate	Standart Deviasi (std)
Spectral Centroid	Standart Deviasi (std)
Spectral Flux	Standart Deviasi (std) dengan nilai rata-rata Flux

4. Implementasi

1. Persiapan Peralatan

Sebelum melakukan perekaman suara harus dilakukan penataan seperti pada gambar berikut ini.



Gambar 9. Desain perangkat perekaman pengukuran energi sinyal audio

PC harus dilengkapi dengan peralatan *multimedia* seperti *sound card*, *speaker active* dan *microphone*. Untuk *microphone* dan *speaker active* bias juga digantikan dengan *head set* lengkap. Sebelum anda memulai praktikum, sebaiknya anda tes dulu, apakah seluruh perangkat *multimedia* anda sudah terintegrasi dengan PC.

2. Menjalankan program klasifikasi suara dengan Matlab

>> klasifikasi_suara

Data hasil perhitungan ekstraksi ciri sinyal audio yang diperoleh dari 10 sampel suara dan direkam pada file “pola.txt”, data yang disimpan adalah data rata-rata dari masing-masing metode ekstraksi ciri.

No	Nama Mahasiswa	Nama File	Short Time Energy	Zero Crossing Rate	Spectral Centroid	Spectral Flux
1	Nesi	nesi1.wav	0.0538388	0.02373	0.0893531	0.13844
2	Ednofri	ednofri4.wav	0.0730546	0.0305071	0.123814	0.473905
3	David	david4.wav	0.0662328	0.023269	0	0.511881
4	Sofyan	sofyan4.wav	0.137326	0.0199186	0	0.732067
5	Gunawan	gunawan4.wav	0.0923871	0.0449648	0.124977	0.483487
6	Herlambang	herlambang4.wav	0.112719	0.0181066	0.181973	0.586998
7	Hendy	hendy4.wav	0.115847	0.0283285	0.116744	0.540119
8	Susapto	susapto1.wav	0.129537	0.0441993	0	0.655706
9	Arif	Arif4.wav	0.0664739	0.0190522	0.173884	0.409927
10	Retno	retno1.wav	0.0866736	0.0283818	0.126805	0.313418

2. Pemanggilan data dari pola.txt dan di rubah ke bentuk matrik

```
>> M = dlmread('pola.txt','')
```

Matrik M yang diperoleh

M =

```

0.0538388    0.02373    0.0893531    0.13844
0.0730546    0.0305071    0.123814    0.473905
0.0662328    0.023269     0          0.511881
0.137326     0.0199186     0          0.732067
0.0923871    0.0449648    0.124977    0.483487
0.112719     0.0181066    0.181973    0.586998
0.115847     0.0283285    0.116744    0.540119
0.129537     0.0441993     0          0.655706
0.0664739    0.0190522    0.173884    0.409927
0.0866736    0.0283818    0.126805    0.313418

```

3. Mengklasifikasi data menggunakan function K-means

```
>> X = kMeansArif(M,2)
```

Hasil yang didapat :

```
maxRow = 10
```

```
maxCol = 4
```

```
c =
```

```
0.0538 0.0237 0.0894 0.1384  
0.0731 0.0305 0.1238 0.4739
```

```
X =
```

0.0538	0.0237	0.0894	0.1384	1.0000
0.0731	0.0305	0.1238	0.4739	2.0000
0.0662	0.0233	0	0.5119	2.0000
0.1373	0.0199	0	0.7321	2.0000
0.0924	0.0450	0.1250	0.4835	2.0000
0.1127	0.0181	0.1820	0.5870	2.0000
0.1158	0.0283	0.1167	0.5401	2.0000
0.1295	0.0442	0	0.6557	2.0000
0.0665	0.0191	0.1739	0.4099	2.0000
0.0867	0.0284	0.1268	0.3134	1.0000



Hasil Clustering

Atau dengan algoritma K-means bawaan Matlab :

```
>> X = kmeans(M,2)
```

```
X = 1 2 2 2 2 2 2 2 2 1
```

Keterangan :

1 = Wanita

2 = Pria

6. Kesimpulan

Dari hasil makalah yang penulis susun, dapat disimpulkan bahwa pengklasifikasian suara berdasarkan jenis kelamin atau gender dapat dilakukan dengan suatu metode ekstraksi ciri sinyal suara berbasis domain waktu dan domain frekuensi. Ekstraksi ciri yang digunakan bisa bermacam-macam, diantaranya adalah dengan Sort Time Energy, Zero Crossing Rate, Spectral Centroid, dan Spectral Flux. Nilai rata-rata dari masing-masing ciri dihitung dengan standart deviasi, untuk memperoleh rata-rata nya, lalu diolah untuk pengklasifikasian.

7. Daftar Pustaka

Ellis. D., "*Design of a Speaker Recognition Code using Matlab*", Department of Computer and Elektrical Engineering, University of Tennessee, 2001.

J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, New York, 1981.

Min N. Do, "*An Automatic Speaker Recognition System*", Audio Visual Communication Laboratory, Swiss Federal Institute of Technology, Lausanne, Switzerland, 2001.

Richards. J.A, *Remote Sensing Digital Image Analysis, An Introduction*, Springer-Verlag Berlin Heidelberg, 1986

Setiawan. A., "*Klasifikasi Suara Berdasarkan Gender Dengan Algoritma K-Means*", Pengenalan Pola, Fakultas MIPA, Universitas Gajah Mada, 2008.

Ulfa. M., Prima K, Titon Dutono, "*Aplikasi Speech Recognition sebagai Fungsi Layanan pada SistemPerbankan*", Seminar SNTE, Jogjakarta 2003.